X-641-71-373
SUPERSEDES
X-641-71-56

X- 65700

HAMILTONIAN FORMULATION OF GUIDING CENTER MOTION

DAVID P. STERN

SEPTEMBER 1971

GSFC — GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

Hamiltonian Formulation of Guiding Center Motion

David P. Stern
Theoretical Studies Branch
Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract

The nonrelativistic guiding center motion of a charged particle in a static magnetic field is derived using the Hamiltonian formalism. By repeated application of first-order canonical perturbation theory the first two adiabatic invariants and their averaged Hamiltonians are derived, including the first-order correction terms. Other features of guiding center theory are also obtained, including lowest order drifts and the flux invariant.

CANONICAL PERTURBATION THEORY

This is an attempt to formulate and solve the nonrelativistic guiding center motion of a charged particle in Hamiltonian terms. In other words, since guiding center motion may be viewed as a perturbed periodic motion, we try to apply to it methods of celestial mechanics (suitably modified) which are designed for motions of this type.

One such method (often associated with the names of Poincaré and Von Zeipel[(1) - (6)]) is the following. One begins by obtaining the zero-order Hamiltonian $H^{(0)}$ — that is, the limiting Hamiltonian for the case of the perturbation tending to zero — and solves the associated Hamilton-Jacobi equation for action-angle variables (J, Ω). This should be possible since it is given that the $\varepsilon \rightarrow 0$ limit is both periodic and soluble.

The Hamilton-Jacobi equation gives the generating function W of a canonical transformation to new variables (p_i, q_i) that include the action-angle pair. In the system described by $H^{(0)}$ alone, all new variables are constants of the motion, except for Ω which is linear in time and increases by unity each period. This transformation is now applied to the full Hamiltonian (with finite ε), so that the new variables vary slowly, except for Ω which is now approximately linear in time.

The prescription next calls for a near-identity canonical transformation

$$(p_i, q_i) \longrightarrow (\mathfrak{p}_i, \mathfrak{q}_i)$$

with transformed action-angle variables (J^*, Ω^*) and with

$$H^{*}(\underline{p}, \underline{q}) = H^{*(0)}(\underline{p}, \underline{q}) + \varepsilon H^{*(1)} + \ldots \tag{1}$$

the new Hamiltonian, such that H^{*} does not contain Ω^{*}. If this is accomplished then J^{*} is a constant of the motion, termed adiabatic invariant in guiding center motion and sometimes called a third integral in celestial mechanics (the terms are not synonymous and their difference will presently be noted). Further solution of the problem by means of H^{*} does not have to contend with (J^{*}, Ω^{*}) and thus involves two variables less than the original problem, corresponding to a reduction by one dimension.

Let the near-identity transformation have a generating function

$$S(\underline{p}, \underline{q}) = \sum p_i q_i + \sum_{k=1} \varepsilon^{k} S^{(k)}(\underline{p}, \underline{q}) \tag{2}$$

Then if the time t is not explicitely involved, the old and new Hamiltonians are equal, giving

$$\sum_{k=0} \varepsilon^{k} H^{(k)}(\underline{p} + \sum \varepsilon^{m} \partial S^{(m)}/\partial \underline{q}, \underline{q}) =$$

$$= \sum_{k=0} \varepsilon^{k} H^{*(k)}(\underline{p}, \underline{q} + \sum \varepsilon^{m} \partial S^{(m)}/\partial \underline{p})$$

Expansion gives, for $O(\varepsilon^{0})$

$$H^{(0)}(\underline{p}, \underline{q}) = H^{*(0)}(\underline{p}, \underline{q}) \tag{3}$$

which is acceptable, as $H^{(0)}$ does not contain Ω, the conjugate J being a constant of the motion. For the terms of order ε one gets

$$H^{(1)} + \sum (\partial S^{(1)}/\partial q_i)(\partial H^{(0)}/\partial p_i) =$$

$$= H^{*(1)} + \sum (\partial S^{(1)}/\partial p_i)(\partial H^{(0)}/\partial q_i) \qquad (4)$$

If this is not to be a partial differential equation for $S^{(1)}$, only one term containing $S^{(1)}$ may be allowed. In celestial mechanics, where the perturbation is "small" and is expressed by terms of various orders in ε added to $H^{(0)}$, this is accomplished by allowing $H^{(0)}$ to contain only J :

$$H^{(0)} = J\omega/2\pi \qquad (\omega = \text{const.}) \qquad (5)$$

Then (4) gives

$$(\omega/2\pi)(\partial S^{(1)}/\partial\Omega) + H^{(1)} = H^{*(1)} \qquad (6)$$

If (as is assumed) any dependence on Ω is via periodic functions, then any function F containing Ω may be resolved into an averaged part

$$\langle F \rangle = \int_0^1 F \, d\Omega \qquad (7)$$

and a "purely periodic" or "oscillating" part, with zero average,

$$(F)_{osc} = F - \langle F \rangle \qquad (8)$$

Now $H^{*(1)}$ has no dependence on Ω while $(\partial S^{(1)}/\partial\Omega)$ is purely periodic — any part of $S^{(1)}$ independent of Ω is eliminated by the differentiation — so

$$H^{*(1)} = \langle H^{(1)} \rangle \tag{9}$$

$$S^{(1)} = -(2\pi/\omega) \int (H^{(1)})_{osc} \, d\Omega \tag{10}$$

which gives the transformation to first order.

In guiding center motion the situation differs somewhat, since the perturbation is not necessarily small, but rather is what is known as <u>slow</u> or <u>adiabatic</u> (this leads to the difference between third integrals and adiabatic invariants). In an adiabatically perturbed system some of the variables (collectively denoted as $\underline{q}_s$ and $\underline{p}_s$) are "slow" and have the property that for any function F appearing in the calculation and depending on them (but not containing them as explicit factors), $\partial F/\partial \underline{q}_s$ or $\partial F/\partial \underline{p}_s$ are of the order of εF . Slow explicit dependence on time is also possible, but will not be discussed here; it has been investigated by Gardner(7), in an article discussed in more detail by Contopoulos(8). Gardner's method differs somewhat from the one described here and further developed by Stern(6), in that it uses a separate canonical transformation for each order in ε.

In adiabatically perturbed motion $H^{(0)}$ may depend on variables other than J , provided that they belong to canonical pairs of which at least one member is "slow". Corresponding terms in (4) are then shifted to higher orders and one again ends up with (6), (9) and (10). Higher-order results differ somewhat from those derived for "small" perturbations, but in this work only the first-order correction terms will be considered.

Another complicating factor that only arises with orders higher than the first is caused by the fact that the functions $S^{(k)}$ representing the transformation depend on "mixed" variables -- old coordinates and new momenta. To express new variables in terms of old ones, or vice versa, further untangling is needed; this may be avoided by using the direct form of near-identity canonical transformations[6] [9] [10]. Here such methods are not required, however, for in calculations correct to the first order it is always permissible -- due to the transformation being a near-identity one -- to replace new variables by old ones in first-order correction terms.

For instance, from the transformation relation

$$p_i \quad = \quad \partial S/\partial q_i \quad = \quad \mathfrak{p}_i \;+\; \sum \varepsilon^k \,\partial S^{(k)}/\partial q_i$$

and from (6) and (9), the adiabatic invariant will be

$$\begin{aligned} J^* &= J - \varepsilon\, \partial S^{(1)}/\partial \Omega + O(\varepsilon^2) \\ &= J + \varepsilon\,(2\pi/\omega)\,(H^{(1)}[\underline{\mathfrak{p}}, \underline{q}])_{osc} + O(\varepsilon^2) \\ &= J + \varepsilon\,(2\pi/\omega)\,(H^{(1)}[\underline{p}, \underline{q}])_{osc} + O(\varepsilon^2) \qquad (11) \end{aligned}$$

Because the first-order correction of J^* is purely periodic, the long-term average of J is conserved to order ε^2. It may be shown from this[6] that the total variation of J over a time period of order ε^{-1} is of order ε, leading to the well-known property of lowest-order adiabatic invariants, namely, that in a system undergoing a finite perturbation their total variation may be made arbitrarily small by stretching out the perturbation over a long enough time.

The arguments of H^* evolve on a time scale one order slower than that of the zero-order gyration of H . If however the motion represented by H^* is also periodic in its lowest order, the preceding procedure may be repeated, leading to additional adiabatic invariants (this was also pointed out by Gardner). A particle in the earth's radiation belt, for instance, may have three independent periodicities, each associated with an adiabatic invariant, and these invariants should be derivable by repeated application of the preceding routine.

Almost all perturbation methods contain a "smallness parameter" $\varepsilon \ll 1$, but in adiabatic perturbation theory ε is somewhat artificial[6]. It may be introduced by rewriting $O(\varepsilon)$ terms (e.g. for some "slow" variable q_{si}) as

$$\partial F/\partial q_{si} \longrightarrow \varepsilon \left(\partial F/\partial(\varepsilon q_{si})\right) \tag{12}$$

but (since no definite numerical value of ε is available) it must be removed by the reverse process before the final result is obtained. The only time ε is used is in grouping terms according to the order to which they belong. In what follows there will generally be no question about the order to which a term belongs, and therefore we may <u>avoid any use of ε altogether</u>, although first-order terms will sometimes be labeled by superscript "(1)" .

PRELIMINARY TRANSFORMATION OF THE HAMILTONIAN

The guiding center motion of a charged particle in a magnetic field is strongly related to the structure of field lines and one would expect this structure to enter somehow into the calculation. This is done via the Euler potentials[11] (α, β), related to the field vector $\underline{B}$ by

$$\underline{B} = \nabla\alpha \times \nabla\beta \qquad (13)$$

Each field line then may be viewed as the intersection of two surfaces

$$\alpha = \text{constant} = \alpha_0$$

$$\beta = \text{constant} = \beta_0$$

and it is consequently labeled by the parameters (α_0, β_0). Associated with (α, β) is a vector potential $\alpha\nabla\beta$ orthogonal to $\underline{B}$, which will be henceforth adopted.

Euler potentials may be introduced into the dynamical problem of charged particle motion by choosing (α, β) as curvilinear coordinates of position. For the third coordinate one may choose s, the distance measured along field lines from some given reference surface. In curl-free magnetic fields the scalar magnetic potential γ would probably be a better choice, but this was not investigated.

Gardner[7] has carried out the first steps in transforming the Hamiltonian and we shall start with his transformation, continuing from there according to the prescription outlined earlier. Our choice of subscripts follows Gardner's, but his use of ε is not adopted.

Let the vector $\underline{x}$ give the cartesian coordinates of a particle of mass m and charge e, and let $\underline{\Pi}$ be the canonical momentum vector conjugate to $\underline{x}$. The non-relativistic Hamiltonian for the particle's motion in a time-independent magnetic field is then

$$H = (1/2m) \left\{ \underline{\Pi} - (e/c) \alpha \nabla \beta \right\}^2 \qquad (14)$$

A canonical transformation to new variables $(\underline{P}, \underline{Q})$, generated by the function

$$F(\underline{P}, \underline{x}) = P_3 \alpha(\underline{x}) + P_1 \beta(\underline{x}) + P_2 s(\underline{x}) \qquad (15)$$

transforms to a new system in which the canonical coordinates Q_i equal (α, β, s). It does not, however, separate the rapid gyration around field lines from other motions. That is achieved by a canonical transformation proposed by Gardner(7), generated by

$$F = P_3 \alpha + P_1 \beta + P_2 s - (c/e) P_3 P_1 \qquad (16)$$

This gives

$$H = (1/2m) \left\{ P_3 \nabla \alpha - (e/c) Q_3 \nabla \beta + P_2 \nabla s \right\}^2 \qquad (17)$$

and

$$\left. \begin{aligned} \alpha &= (c/e) P_1 + Q_3 \\ \beta &= Q_1 + (c/e) P_3 \\ s &= Q_2 \end{aligned} \right\} \qquad (18)$$

In a homogeneous field with $\underline{B} = \hat{z} B$ ($\hat{z}$ unit vector) this is readily solved(11) by choosing

$$\alpha = x \qquad \beta = B\,y \qquad s = z \tag{19}$$

leading to

$$H = (1/2m)\left[P_3^{\,2} + (e/c)^2 B^2 Q_3^{\,2} + P_2^{\,2}\right] \tag{20}$$

The variables (P_3, Q_3) appear here in the same form as the canonical variables in the Hamiltonian of the simple harmonic oscillator; therefore, they vary periodically and represent the gyration. On the other hand, by (18), $c\,P_1/e$ and Q_1 may be identified as the (constant) Euler potentials of the guiding field line, while P_2 represents the momentum component parallel to the field and is also constant.

Unfortunately, this solution cannot be immediately extended to the general nearly-homogeneous field, because the vectors $(\nabla\alpha, \nabla\beta, \nabla s)$ are usually not orthogonal. The non-orthogonality of $(\nabla\alpha, \nabla\beta)$ can be remedied by adding to the generating function F a term $\frac{1}{2}\lambda P_3^{\,2}$, where

$$\lambda = (c/e)\,(\nabla\alpha\cdot\nabla\beta)/\,(\nabla\beta\,)^2 \tag{21}$$

The non-orthogonality of ∇s to the two other vectors is harder to correct. One might hope that some scalar s' replacing s would have the required property, but this is not so, except for special cases (e.g. curl-free fields, where the scalar potential γ may serve as s'). For if s' existed such that $\nabla s'\cdot\nabla\alpha$ and $\nabla s'\cdot\nabla\beta$ both vanished, then a scalar μ must exist such that

$$\underline{B} = \mu\,\nabla s'$$

implying

$$\underline{B}\cdot(\nabla\times\underline{B}) = 0$$

which is not always satisfied.

However, if the field varies slowly in space, so that quantities such as $\nabla\nabla\alpha$ and $\nabla\nabla\beta$ are $O(\varepsilon)$, it is possible to define a scalar σ such that $\nabla\sigma$ is orthogonal to $\nabla\alpha$ and $\nabla\beta$ within $O(\varepsilon)$ correction terms. We take

$$\sigma = s - \alpha a - \beta b \tag{22}$$

where

$$\begin{aligned} a &= B^{-2}\ \underline{B}\cdot(\nabla s \times \nabla\beta) \\ b &= B^{-2}\ \underline{B}\cdot(\nabla\alpha \times \nabla s) \end{aligned} \tag{23}$$

giving ($\hat{\underline{B}} = \underline{B}/B$)

$$\hat{\underline{B}} = \nabla s - a\nabla\alpha - b\nabla\beta \tag{24}$$

and hence

$$\begin{aligned} \nabla\sigma &= \hat{\underline{B}} - \alpha\nabla a - \beta\nabla b \\ &= \hat{\underline{B}} + O(\varepsilon) \end{aligned} \tag{25}$$

In this calculation the field $\underline{B}$ and associated quantities such as $\nabla\alpha$, $\nabla\beta$, a, b and λ are all $O(1)$, while quantities derived from them by differentiation — $\nabla\nabla\alpha$, ∇a, $\nabla\lambda$ etc. — are all $O(\varepsilon)$. However, α and β appearing explicitely in eq. (25) are not of order ε^{-1} but also $O(1)$. The difference between them and $O(1)$ field quantities such as $\underline{B}$ is that field quantities depend slowly (or adiabatically) on position variables, whereas α and β depend on such variables in a "regular" (non-slow) fashion. This is evident from (18) for the variables generated by (16) and also holds in what follows.

To accomodate the non-orthogonality of the basic vectors we let the initial canonical transformation be generated by

$$F(\underline{P}, \underline{x}) = \alpha P_3 + \beta P_1 + \sigma P_2 - (c/e) P_3 P_1 + \tfrac{1}{2} \lambda P_3^{\,2} \qquad (26)$$

Application of the transformation equations gives

$$\alpha = (c/e) P_1 + (Q_3 + \lambda P_3) \qquad (27)$$

$$\beta = Q_1 + (c/e) P_3 \qquad (28)$$

$$\sigma = Q_2 \qquad (29)$$

Also, the canonical momentum vector is (gradient taken in $\underline{x}$ space)

$$\underline{\pi} = \nabla F = P_3 \nabla\alpha + P_1 \nabla\beta + P_2 \nabla\sigma - \tfrac{1}{2} P_3^{\,2} \nabla\lambda$$

Substituting everything in (14) then yields

$$H = (1/2m) \left\{ P_3 \nabla\alpha - (e/c)(Q_3 + \lambda P_3) \nabla\beta + P_2 \nabla\sigma - \tfrac{1}{2} P_3^{\,2} \nabla\lambda \right\}^2 \qquad (30)$$

The vector being squared is the (non-canonical) momentum $m\,\underline{v}$ of the particle and the first two terms in it may be regarded as the zero-order component $m\,\underline{v}_{\perp}^{(0)}$ of this vector orthogonal to the magnetic field. If one defines a vector $\underline{T}$ orthogonal to both $\underline{B}$ and $\nabla\beta$

$$\underline{T} = \nabla\alpha - (e/c)\lambda \nabla\beta \qquad (31)$$

then this part may be written in terms of orthogonal vectors as

$$m\,\underline{v}_{\perp}^{(0)} = P_3 \underline{T} - (e/c) Q_3 \nabla\beta \qquad (32)$$

THE GUIDING CENTER EXPANSION

The magnitude of $\underline{T}$ is

$$|\underline{T}| = B/|\nabla\beta| \tag{33}$$

which with (30) and (32) gives

$$H = (1/2m)\left\{P_3^{\,2}\,B^2/(\nabla\beta)^2 + (e/c)^2\,Q_3^{\,2}(\nabla\beta)^2 + P_2^{\,2}(\nabla\sigma)^2\right\} + O(\varepsilon) \tag{34}$$

This resembles (20), but there exists an important difference: the variables (P_3, Q_3) appear not only as explicit factors but also by (27)-(29), in the variables (α, β, σ) giving the locations at which B^2 and $(\nabla\beta)^2$ are to be evaluated. To eliminate (P_3, Q_3) from these variables one must expand any function of position around the "guiding center" point, denoted by subscript "c" and defined by

$$(\alpha_c, \beta_c, \sigma_c) = (cP_1/e, Q_1, Q_2) \tag{35}$$

By (27)-(30), the guiding center expansion for any function f of position may be written

$$f(\alpha, \beta, \sigma) = (\exp D)\ f(\alpha_c, \beta_c, \sigma_c)$$

where D is the operator

$$D = (Q_3 + \lambda P_3)(\partial/\partial\alpha_c) + (c\,P_3/e)(\partial/\partial\beta_c) \tag{36}$$

and where $(\exp D)$ is defined by the Taylor series of the exponential function. In what follows, all functions depending on position will be assumed to be evaluated at the guiding center, unless otherwise is stated.

To understand the significance of D , let us define (in guiding center variables, subscript omitted) the operator

$$\nabla_\perp = \nabla\alpha(\partial/\partial\alpha) + \nabla\beta(\partial/\partial\beta) \tag{37}$$

In addition, let the <u>angular frequency of gyration</u> be defined as

$$\omega = Be/mc \tag{38}$$

and the <u>gyration radius</u> as the vector

$$\underline{\rho} = \omega^{-1}(\hat{\underline{B}} \times \underline{v}_\perp^{(0)}) \tag{39}$$

with $\underline{v}_\perp^{(0)}$ defined as in (32) but evaluated at the guiding center. One then finds (it is best not to simplify the cross products below)

$$(\underline{\rho}\cdot\nabla_\perp) = -(\omega B)^{-1}\, \underline{v}_\perp^{(0)}\cdot\Big\{(\underline{B}\times\nabla\alpha)(\partial/\partial\alpha) + (\underline{B}\times\nabla\beta)(\partial/\partial\beta)\Big\} = D \tag{40}$$

Two comments may be made here. First, in order for the expansion to be valid we require

$$D \ll 1$$

i.e.

$$(\underline{\rho}\cdot\nabla_\perp) \ll 1 \tag{41}$$

In other words, $\underline{\rho}$ is to be much smaller than the scale on which the field variables change. This is known as <u>Alfvén's criterion</u> for the validity of the guiding center approximation and provides a more precise formulation of the statement made earlier that "the field varies slowly in space."

Secondly, the difference between $\underline{\rho}\cdot\nabla_{\perp}$ and $\underline{\rho}\cdot\nabla$ is small, since $\underline{B}$ and $\nabla\sigma$ are almost parallel; it is given by

$$(\underline{\rho}\cdot\nabla\sigma)\,\partial/\partial\sigma = \omega^{-1}(\nabla\sigma\times\hat{\underline{B}})\cdot\underline{v}_{\perp}^{(0)}\,\partial/\partial\sigma$$

$$= O(\varepsilon)\,\partial/\partial\sigma \qquad (42)$$

In the present calculation we only retain terms of order ε and therefore in such terms the difference between $\underline{\rho}\cdot\nabla_{\perp}$ and $\underline{\rho}\cdot\nabla$ may be ignored.

Application of the expansion to the Hamiltonian (30) then gives

$$H = (1/2m)\left\{m\,\underline{v}_{\perp}^{(0)} + P_2\nabla\sigma + m\,\underline{v}^{(1)}\right\}^2 + O(\varepsilon^2) \qquad (43)$$

where

$$m\,\underline{v}^{(1)} = m\,\underline{\rho}\cdot\nabla\underline{v}_{\perp}^{(0)} + P_2\,\underline{\rho}\cdot\nabla\nabla\sigma - \tfrac{1}{2}P_3^{\,2}\nabla\lambda = O(\varepsilon) \qquad (44)$$

and where all variables are defined at the guiding center.

ZERO ORDER ACTION - ANGLE VARIABLES

We now apply to the system the transformation derived from the Hamilton-Jacobi equation in the limit $\varepsilon \rightarrow 0$. In this limit both B^2 and $(\nabla\beta)^2$ are constant and the Hamilton-Jacobi equation derived from (34) separates into two equations, representing motion parallel and perpendicular to the field. Of these the former gives constant motion along the field's direction, while the latter resembles the Hamilton-Jacobi equation of a harmonic oscillator(12)(13) . Of course, when we actually apply this transformation to the given system, B^2 and $(\nabla\beta)^2$ must be allowed to vary slowly again.

If the new variables are denoted by $(\underline{p}, \underline{q})$, with action-angle variables

$$(p_3, q_3) = (J, \Omega) \tag{45}$$

then the generating function of the transformation is

$$W(\underline{p}, \underline{Q}) = \int \left[(J\delta\, e/\pi c) - (e\delta/c)^2 Q_3^{\,2} \right]^{\frac{1}{2}} dQ_3 + p_1 Q_1 + p_2 Q_2 \tag{46}$$

where

$$\delta = (\nabla\beta)^2 / B \tag{47}$$

and where all variables are evaluated at the "mixed coordinate guiding center" $\underline{r}'_c$, defined by

$$(\alpha'_c , \beta'_c , \sigma'_c) = (c\, p_1 / e , Q_1 , Q_2) \tag{48}$$

We then have

$$P_3 = \partial W / \partial Q_3 = \left[(J\delta\, e/\pi c) - (e\delta/c)^2 Q_3^{\,2} \right]^{\frac{1}{2}} \tag{49}$$

from which the first part of (34) is

$$(1/2m) \left[P_3^{\,2} B^2 / (\nabla\beta)^2 + (e/c)^2 Q_3^{\,2} (\nabla\beta)^2 \right] = J\omega / 2\pi \tag{50}$$

Strictly speaking, this may not be directly substituted in the Hamiltonian since ω and other quantities are here defined in mixed variables. However, we have for such quantities

$$\omega(\underline{r}'_c) = \omega(\underline{r}_c) + (c/e)\ (P_1 - p_1)\ (\partial\omega/\partial\alpha) + \ldots$$

$$= \omega(\underline{r}_c) + O(\varepsilon^2) \qquad (51)$$

since $\partial\omega/\partial\alpha$, $(P_1 - p_1)$ and similar quantities are all $O(\varepsilon)$. Because the present calculation ignores second-order terms, the difference between "mixed" and "true" guiding center variables (either "new" or "old") in slowly varying functions will be neglected.

One also has

$$\Omega = \partial W/\partial J \qquad (52)$$

from which

$$Q_3 = (Jc/\pi\,\delta\,e)^{\frac{1}{2}}\ \sin\,(2\pi\Omega) \qquad (53)$$

$$P_3 = (J\,\delta\,e/\pi\,c)^{\frac{1}{2}}\ \cos\,(2\pi\Omega) \qquad (54)$$

Because only zero and first order terms are retained all other variables may be assumed to transform identically, <u>except</u> for P_2 , which appears explicitely in the zero-order Hamiltonian. For P_2 one finds

$$P_2 = \partial W/\partial Q_2 = p_2 + (\partial W/\partial\delta)(\partial\delta/\partial\sigma) + O(\varepsilon^2)$$

$$= p_2 + \Delta p_2 + O(\varepsilon^2) \qquad (55)$$

It may be shown that the integral $\partial W/\partial\delta$ reduces to

$$\partial W/\partial\delta = Q_3\,P_3\,/\,2\,\delta = (J/4\pi\,\delta)\ \sin\,(4\pi\Omega) \qquad (56)$$

which allows Δp_2 to be computed. The new Hamiltonian may thus be written

$$H = H^{(0)} + H^{(1)} + O(\varepsilon^2) \tag{57}$$

where the zero-order part is

$$H^{(0)} = J\omega/2\pi + (\nabla\sigma)^2\, p_2{}^2/2m \tag{58}$$

and the first-order correction is

$$\begin{aligned} H^{(1)} = {} & m(\underline{v}_\perp^{(0)} \cdot \underline{v}^{(1)}) + p_2 \nabla\sigma \cdot (\underline{v}_\perp^{(0)} + \underline{v}^{(1)}) \\ & + p_2\, \Delta p_2 / 2m \end{aligned} \tag{59}$$

DRIFTS

From the transformed Hamiltonian we can derive the slow variation of the guiding center coordinates (α_c, β_c), which may be conveniently redefined as $(cp_1/e, q_1)$. This represents an $O(\varepsilon)$ difference from (35), but since only the lowest order of the motion will be derived, this difference may be ignored.

One can always assign to the guiding center a velocity $\underline{V}$ satisfying (subscript "c" omitted)

$$\begin{aligned} \dot{\alpha} &= (\underline{V} \cdot \nabla\alpha) + \partial\alpha/\partial t \\ \dot{\beta} &= (\underline{V} \cdot \nabla\beta) + \partial\beta/\partial t \end{aligned} \tag{60}$$

In the present case $\partial/\partial t$ vanishes and Hamilton's equations give

$$\dot{\alpha} = c\,\dot{p}_1/e = -(c/e)\;\partial H/\partial q_1 = -(c/e)\;\partial H/\partial\beta$$
$$\dot{\beta} = \dot{q}_1 = \partial H/\partial p_1 = (c/e)\;\partial H/\partial\alpha \tag{61}$$

Hence

$$\begin{aligned}\underline{V}\times\underline{B} &= \underline{V}\times(\nabla\alpha\times\nabla\beta)\\ &= \nabla\alpha\;(\underline{V}\cdot\nabla\beta) - \nabla\beta\;(\underline{V}\cdot\nabla\alpha)\\ &= \nabla\alpha\;\dot{\beta} - \nabla\beta\;\dot{\alpha}\\ &= (c/e)\;\nabla_{\perp}H \end{aligned}\tag{62}$$

The component of $\underline{V}$ parallel to $\underline{B}$ is not determined by (60) and by the equation that follow it and will therefore be set equal to zero, since we only are interested in motion perpendicular to field lines. Forming the vector product with $\underline{B}$ then gives

$$\underline{V} = (c/eB^2)\;\underline{B}\times\nabla_{\perp}H \tag{63}$$

Because higher order corrections are not derived, we may again ignore the difference between $\nabla_{\perp}$ and ∇. Using (58) we find that ∇H consists of two parts, each of which contributes to the "drift motion" represented by $\underline{V}$. The first part contributes the so-called <u>gradient drift</u>

$$\begin{aligned}\underline{V}_g &= (c/eB^2)\;J/2\pi\,(\underline{B}\times\nabla\omega)\\ &= (c/eB^2)\;J\omega/2\pi\;(\hat{\underline{B}}\times\nabla B)\end{aligned}\tag{64}$$

Because of (50)

$$J\omega/2\pi = \tfrac{1}{2} m v_\perp^2 + O(\varepsilon) \qquad (65)$$

from which

$$\underline{V}_g = (v_\perp^2 /2\omega B) \; \hat{\underline{B}} \times \nabla B \qquad (66)$$

which is the customary expression for this drift. The second term contributes the so-called curvature drift

$$\underline{V}_{curv} = (c\, p_2{}^2/2emB^2) \quad \underline{B} \times \nabla(\nabla\sigma)^2 \qquad (67)$$

By (25)

$$\begin{aligned} \tfrac{1}{2} \nabla(\nabla\sigma)^2 &= \nabla\sigma \cdot \nabla\nabla\sigma \\ &= \left[\hat{\underline{B}} + O(\varepsilon) \right] \cdot \left[\nabla\hat{\underline{B}} - \nabla\alpha\nabla a - \nabla\beta\nabla b + O(\varepsilon^2) \right] \\ &= \hat{\underline{B}} \cdot \nabla\hat{\underline{B}} + O(\varepsilon^2) \end{aligned} \qquad (68)$$

with the last equality following the vanishing of $\underline{B}\cdot\nabla\alpha$ and $\underline{B}\cdot\nabla\beta$. To lowest order one may also replace p_2 with $m\, v_{/\!/}$, where $v_{/\!/}$ is the velocity component parallel to the field, giving the drift in its usual form as

$$\underline{V}_{curv} = (v_{/\!/}^2 /\omega) \; \hat{\underline{B}} \times (\hat{\underline{B}} \cdot \nabla\hat{\underline{B}}) \qquad (69)$$

If the $\partial/\partial t$ terms of (60) had been retained, then (63) would have contained an additional term

$$\underline{V}_F = (c/eB^2) \; \underline{B} \times \left\{ \nabla\beta(\partial\alpha/\partial t) - \nabla\alpha(\partial\beta/\partial t) \right\} \qquad (70)$$

representing the velocity with which field lines appear to move due to $\partial\alpha/\partial t$ and $\partial\beta/\partial t$ [11][14][15]. While this certainly does not represent all effects of time dependence, a term of this form may be expected to appear when one deals with a time-dependent field.

THE FIRST ORDER INVARIANT

By (11) and (59) the adiabatic invariant associated with gyration around field lines is

$$J^* = J + (2\pi/\omega m)\Big\{ m^2 (\underline{v}_\perp^{(0)} \cdot \underline{v}^{(1)})_{osc} + p_2 \Delta p_2 + m\, p_2 \nabla\sigma \cdot (\underline{v}_\perp^{(0)} + \underline{v}^{(1)})_{osc} \Big\} + 0(\varepsilon^2) \qquad (71)$$

where ε is not explicitely written since it is already included in the present definition of $H^{(1)}$ and where the subscript "osc" is omitted from the term containing Δp_2 since, by (55) and (56), this term is purely periodic with zero average.

The above result is not in useful form, since J is not directly observed. It is better to eliminate J in favor of $v_\perp^2$, the square of the total velocity component perpendicular to the field, the field's direction being chosen as that existing at the guiding center. By (32), (50) and (51)

$$J = \pi m (\underline{v}_\perp^{(0)})^2/\omega + 0(\varepsilon^2) \qquad (72)$$

In $0(\varepsilon)$ terms this may be used with the superscript zero omitted, as in (65). In the first term of (71), however, J must be expressed accurately to order ε.

Let

$$\underline{v}_\perp = \underline{v}_\perp^{(0)} + \Delta \underline{v}_\perp \qquad (73)$$

Then with $O(\varepsilon)$ accuracy

$$v_\perp^2 = (\underline{v}_\perp^{(0)})^2 + 2(\underline{v}_\perp^{(0)} \cdot \Delta \underline{v}_\perp)$$

$$= J\omega/2\pi + 2(\underline{v}_\perp^{(0)} \cdot \Delta \underline{v}_\perp) \qquad (74)$$

From (43), (25) and (35), to lowest order

$$\Delta \underline{v}_\perp = \underline{v}_\perp^{(1)} + m^{-1} p_2 \nabla_\perp \sigma$$

$$= \underline{v}_\perp^{(1)} - m^{-1} p_2 \left[(c/e)\, p_1 \nabla_\perp a + q_1 \nabla_\perp b \right] \qquad (75)$$

Before substituting this in (71) one may note that to the order of that equation

$$m\, p_2 \nabla\sigma \cdot (\underline{v}_\perp^{(0)} + \underline{v}^{(1)}) = m\, p_2 (\nabla_\perp\sigma \cdot \underline{v}_\perp^{(0)}) + m\, p_2 (\hat{\underline{B}} \cdot \underline{v}^{(1)}) \qquad (76)$$

The combination of (71) and (75) therefore gives

$$J = (\pi m v_\perp^2/\omega) - (2\pi/\omega m) \Big\{ m^2 \langle \underline{v}_\perp^{(0)} \cdot \underline{v}_\perp^{(1)} \rangle +$$

$$+ m\, p_2 \langle \nabla_\perp\sigma \cdot \underline{v}_\perp^{(0)} \rangle - m\, p_2 (\hat{\underline{B}} \cdot \underline{v}^{(1)})_{osc} - p_2 \Delta p_2 \Big\}$$

$$+ O(\varepsilon^2) \qquad (77)$$

The averaged terms vanish. In evaluating them we may discard all terms that are linear or cubic in components of $\underline{v}_\perp^{(0)}$ — including such factors as Q_3, P_3 and $\underline{\rho}$ — since such terms average to zero. By (75), the term in (77) involving $\nabla_\perp \sigma$ is linear in $\underline{v}_\perp^{(0)}$ and vanishes upon averaging. In deriving the term preceding it we may omit all terms of $\underline{v}^{(1)}$ that are quadratic in oscillating quantities, which by (44) only leaves

$$m^2 \left\langle \underline{v}_\perp^{(0)} \cdot \underline{v}_\perp^{(1)} \right\rangle = m\, p_2 \left\langle \underline{v}_\perp^{(0)} \underline{\rho} : \nabla\nabla\sigma \right\rangle \tag{78}$$

Now it may be shown from (31) that

$$\underline{B} \times \nabla\beta = -\underline{T}\,(\nabla\beta)^2 \tag{79}$$

$$\underline{B} \times \underline{T} = \nabla\beta\; B^2/(\nabla\beta)^2 \tag{80}$$

Combining these with (32) and (39) gives

$$\underline{\rho} = (1/m\omega B)\left\{\nabla\beta\, P_3\, B^2/(\nabla\beta)^2 + \underline{T}\,(e/c)\, Q_3\, (\nabla\beta)^2\right\} \tag{81}$$

By (56) the average of $P_3\, Q_3$ vanishes, while (compare eq. 50)

$$(e/c)^2 \left\langle Q_3^{\,2} \right\rangle (\nabla\beta)^2 = \left\langle P_3^{\,2} \right\rangle B^2/(\nabla\beta)^2$$

$$= J\omega m/2\pi \tag{82}$$

Therefore, from (32) and (81)

$$\left\langle \underline{v}_\perp^{(0)} \underline{\rho} \right\rangle = (J/2\pi m B)\left[\underline{T}\,(\nabla\beta) - (\nabla\beta)\,\underline{T}\right] \tag{83}$$

This is an antisymmetric dyadic and its double scalar product with the symmetric $\nabla\nabla\sigma$ vanishes. Thus if one ignores correction terms purely periodic in Ω , $v_\perp^2/B$ is invariant even to order ε, provided B is evaluated at the guiding center. This result has already been noted[16] .

To derive the first remaining $O(\varepsilon)$ term in (77) one requires the oscillating part of $\underline{v}^{(1)}$. From (44)

$$m\, p_2\ \hat{\underline{B}} \cdot \underline{v}^{(1)}_{osc} = p_2^{\,2}\ \hat{\underline{B}}\underline{\rho} : \nabla\nabla\sigma \quad + \qquad\qquad$$

$$+\ p_2 \left(m\ \hat{\underline{B}}\underline{\rho} : \nabla \underline{v}_{\perp}^{(0)} - \tfrac{1}{2} P_3^{\,2}\, (\hat{\underline{B}} \cdot \nabla\lambda) \right)_{osc} \tag{84}$$

Actually one must excercise care with the middle term appearing on the right. When $\underline{\rho} \cdot \nabla\, \underline{v}_{\perp}^{(0)}$ was originally derived for this term, the ∇ operator did not act on P_3 and Q_3 but only on the factors preceding them:

$$m\ \hat{\underline{B}}\underline{\rho} : \nabla \underline{v}_{\perp}^{(0)} = (\hat{\underline{B}}\underline{\rho} : \nabla\nabla\alpha)\, P_3 - (e/c)(\hat{\underline{B}}\underline{\rho} : \nabla\nabla\beta)(Q_3 + \lambda P_3) \tag{85}$$

(the term containing $\nabla\lambda$ has a vanishing factor $\hat{\underline{B}} \cdot \nabla\beta$). At the present stage, however, Q_3 and P_3 are regarded as functions of (J, Ω) through the relations (53) and (54). These relations also introduce a dependence on guiding center variables, through the function δ which they contain. Thus when $\underline{v}_{\perp}^{(0)}$ is viewed as a function of (J, Ω), $\nabla \underline{v}_{\perp}^{(0)}$ contains extra terms with

$$\begin{aligned} \nabla Q_3 &= -\,(Q_3/2\delta)\ \nabla\delta \\ \nabla P_3 &= (P_3/2\delta)\ \nabla\delta \end{aligned} \tag{86}$$

In equation (84), however, this redefinition of $\nabla\, \underline{v}_{\perp}^{(0)}$ makes no difference, since the extra terms are there accompanied by the vanishing factors $(\hat{\underline{B}} \cdot \nabla\alpha)$ and $(\hat{\underline{B}} \cdot \nabla\beta)$.

By equations (25), within higher orders, $(\hat{\underline{B}} \cdot \nabla \lambda)$ equals $\partial \lambda / \partial \sigma$. Also, by (54)

$$(p_3^{\,2})_{osc} = (e/c)\,(J\delta/2\pi)\cos(4\pi\Omega) \qquad (87)$$

Using preceding results and expressing Δp_2 through (6) gives as the final form of the invariant

$$J^* = \pi m v_\perp^2/\omega + (2\pi/\omega)\Big[\, p_2^{\,2}\,(\hat{\underline{B}}\,\underline{\rho} : \nabla\nabla\sigma) + m\,p_2\,(\hat{\underline{B}}\,\underline{\rho} : \nabla\underline{v}_\perp^{(0)})_{osc} - (e/c)\,p_2\,(J\delta/4\pi)(\partial\lambda/\partial\sigma)\cos(4\pi\Omega) + p_2\,(J/4\pi\delta)(\partial\delta/\partial\sigma)\sin(4\pi\Omega)\Big] + O(\varepsilon^2) \qquad (88)$$

The various first-order terms may be transformed to the forms used by Northrop and by others[15] . This yields all the usual first-order terms, plus one term which has not been accounted for and which may be due to a different definition of the guiding center, due to an error or else may vanish by some identity.

THE TRANSFORMED HAMILTONIAN

The transformation which replaces (J, Ω) with (J^*, Ω^*) affects the remaining variables only in their second order. If the generating function of this transformation is of the form (2), then (for instance) the variable p_2^* replacing p_2 is (compare equation 11)

$$p_2^* = p_2 - \varepsilon\,\partial S^{(1)}/\partial q_2 + O(\varepsilon^2) \qquad (89)$$

and since q_2 is "slow", the second term is also of order ε^2 . We will therefore not change the notation of such variables but assume that they transform identically. The new Hamiltonian H^* then has a zero-order part resembling (58) :

$$H^* \;=\; J^*\omega/2\pi \;+\; (\nabla\sigma)^2\, p_2{}^2/2m \;+\; H^{*(1)} \;+\; O(\varepsilon^2) \tag{90}$$

The first-order part is given by (9) and (59) as

$$H^{*(1)} \;=\; \left\langle \underline{v}_{\perp}^{(0)} \cdot (m\,\underline{v}^{(1)} + p_2 \nabla_{\perp}\sigma) \right\rangle \;+\; p_2 \left\langle \underline{\hat{B}} \cdot \underline{v}^{(1)} \right\rangle \;+\; m^{-1} p_2 \left\langle \Delta p_2 \right\rangle \tag{91}$$

but as was shown in the last section, only the middle term here does not vanish. The factor $(\nabla\sigma)^2$ will be kept in the zero order part, even though it could be resolved into terms of zeroth order (= unity) and of first order. In principle, the first-order part of this expression could cause trouble, since it contains terms proportional to α and β , and one of these may be angle-like and grow without bounds. Formally one can resolve this by switching to a new (α, β) system when the angle-like variable has completed a full circuit; in practice this ambiguity may be expected to disappear when results are expressed in observable quantities and in any case, at a later stage $(\nabla\sigma)^2$ will be transformed away.

Using (44) one finds, to the order of the approximation

$$H^{*(1)} \;=\; p_2 \left\langle \underline{\hat{B}}\, \underline{\rho} : \nabla \underline{v}_{\perp}^{(0)} \right\rangle \;-\; \tfrac{1}{2}\, m^{-1} p_2 \left\langle p_3{}^2 \right\rangle \partial\lambda/\partial\sigma \tag{92}$$

where the contribution of the middle term is omitted since it is linear in oscillating quantities.

The evaluation of the first term requires the vector identity[17]

$$\underline{A}\,\underline{B} : \nabla \underline{C} \quad = \quad \underline{B}\,\underline{A} : \nabla \underline{C} \quad + \quad (\underline{B} \times \underline{A}) \cdot (\nabla \times \underline{C}) \tag{93}$$

(the notation of the preceding equation and of the two that follow is independent of the one used in the rest of this work). Now let $\underline{A}$ and $\underline{C}$ be two orthogonal vectors. Then

$$\nabla(\underline{A}\cdot\underline{C}) = 0 = \underline{A}\cdot\nabla\underline{C} + \underline{C}\cdot\nabla\underline{A} + \underline{A}\times(\nabla\times\underline{C}) + \underline{C}\times(\nabla\times\underline{A}) \tag{94}$$

Performing the scalar multiplication of (94) with an arbitrary vector $\underline{B}$ and applying (93) then gives

$$\underline{A}\,\underline{B} : \nabla\underline{C} \quad = \quad - \,\underline{C}\,\underline{B} : \nabla\underline{A} \tag{95}$$

By (39) the magnetic field and $\underline{\varrho}$ are orthogonal; therefore, using (83) and (93)

$$\begin{aligned}
\left\langle \underline{B}\,\underline{\varrho} : \nabla_{\perp}\underline{v}^{(0)} \right\rangle \quad &= - \left\langle \underline{v}_{\perp}^{(0)}\,\underline{\varrho} \right\rangle : \nabla\underline{B} \\
&= - \left(J^{*} / 2\pi m B \right) \left[\underline{T}\,\nabla\beta - \nabla\beta\,\underline{T} \right] : \nabla\underline{B} \\
&= - \left(J^{*} / 2\pi m B \right) (\nabla\beta \times \underline{T}) \cdot (\nabla \times \underline{B})
\end{aligned} \tag{96}$$

Here one must use J^{*} as variable, since we are dealing with a term of H^{*}. By (31)

$$\underline{T} \times \nabla\beta \quad = \quad \underline{B} \tag{97}$$

so

$$\langle \underline{B}\,\underline{\rho} : \nabla \underline{v}_{\perp}^{(0)} \rangle = (J^*/ 2\pi m)\ \hat{\underline{B}} \cdot (\nabla \times \underline{B}) \qquad (98)$$

To lowest order J^* equals J, and by (82) one may therefore substitute in the last term of (92)

$$\langle p_3{}^2 \rangle = J^* \delta\, e/2\pi c + O(\varepsilon) \qquad (99)$$

giving as final result

$$H^{*(1)} = (p_2\, J^*/2\pi mB)\ \hat{\underline{B}} \circ (\nabla \times \underline{B}) - (p_2\, J^* \delta\, e/4\pi mc)\ \partial\lambda/\partial\sigma \qquad (100)$$

The possible applications of H^* include the associated Liouville equation, which should lead to a gyration-averaged form of the kinetic equation for a collisionless plasma, of the type first derived by Chew, Goldberger and Low[18].

THE SECOND PERIODICITY — THE LIMIT $\varepsilon = 0$

Consider the limiting case of (90) in which higher order terms and the dependence of on (p_1, q_1) may both be neglected. The Hamiltonian

$$H^{*(0)} = J^* \omega (q_2)/2 + (\nabla \sigma)^2\, p_2{}^2/2m \qquad (101)$$

may then be viewed as describing the motion of a particle in a potential proportional to $\omega(q_2)$. If the dependence of ω on q_2 (i.e. on σ of the guiding center) has the form of a potential well deep enough to trap the particle, the motion will be periodic and one can again derive action-angle variables and an adiabatic invariant.

The definition of σ is not unique, since there exists a large arbitrary element in the choice of (α, β, s), but the dependence of field quantities on it always describes the variation along a field line, since the other guiding-center variables already specify the field line itself. If the field intensity B (and therefore ω) is large at two separated points on a field line and lesser between them, a "potential well geometry" will exist and under suitable conditions the particle will be trapped in it. These conditions involve the energy integral E of (101) : since the second term in that equation is non-negative, if the particle starts between two values of q_2 for which $J^*\omega(q_2)/2\pi$ exceeds E , it will be trapped.

Assuming that periodicity exists, we now seek the generating function $W(J_2, q_2)$ to new variables (J_2, Ω_2) that are action-angle variables of (101). Since this is the limit $\varepsilon = 0$, both p_1 and q_1 are to be considered as constant parameters that do not vary in time; independently of ε , J^* is always a constant at this stage and may be removed from consideration as a canonical variable.

If $T(p_1, q_1, J^*)$ is the period of this motion, Ω_2 must increase by unity during each period and therefore the new Hamiltonian must depend on J_2 through a term J_2/T . To derive T one must solve the zero-order motion in the given potential well, something that can often be only done numerically or approximately. This is one point of difference between this periodic motion and the gyration around field lines discussed earlier: with the gyration the basic frequency requires only knowledge of the local magnetic field, while here a more complicated solution of the zero-order periodicity is required.

Assuming T to be known, the Hamilton-Jacobi equation becomes

$$J^*\omega/2\pi \; + \; (\nabla\sigma)^2(\partial W/\partial q_2)^2/2m \; = \; J_2 \,/\, T(p_1, q_1, J^*) \; + \; F \qquad (102)$$

where F is a yet undetermined function of the slow variables (not related to eq. 26). From this

$$W \; = \; (2m)^{\frac{1}{2}} \int \left[J_2/T \; + \; F \; - \; J^*\omega/2\pi \right]^{\frac{1}{2}} \left[dq_2 \,/(\nabla\sigma)^2 \right] \qquad (103)$$

giving

$$(\nabla\sigma)^2 \, p_2{}^2/2m \; = \; J_2/T \; + \; F \; - \; J^*\omega/2\pi \qquad (104)$$

which confirms the form of the new Hamiltonian as the one given on the right of (102). Also

$$\begin{aligned} \Omega_2 \; &= \; \partial W/\partial J_2 \\ &= \; (m/2)^{\frac{1}{2}} \, T^{-1} \int \left[J_2/T \; + \; F \; - \; J^*\omega/2\pi \right]^{-\frac{1}{2}} \left[dq_2 \,/(\nabla\sigma)^2 \right] \end{aligned} \qquad (105)$$

The significance of this is easily seen. Since $J^*\omega/2\pi$ is the energy associated with the component of the motion orthogonal to the field, we have (at least to zero order)

$$v_{\parallel} \; = \; p_2 \,/\, m$$

Also, since along a field line (α, β) do not change

$$dq_2 \; = \; d\sigma \; = \; ds \; + \; O(\varepsilon \, ds) \qquad (106)$$

Hence, if τ is the time of motion evaluated using $v_{\parallel}$ alone

$$\Omega_2 \cong T^{-1} \int (ds/v_{\parallel}) = \tau / T \tag{107}$$

In order to express T we use the fact that Ω_2 increases by unity each period. Integrating (105) over one full period then gives

$$T = m \oint (dq_2/p_2) \cong \oint (ds/v_{\parallel}) \tag{108}$$

It is interesting to note that the familiar formula

$$J_2 = \oint p_2 \, dq_2 \tag{109}$$

is not obtained as part of the calculation. It enters as follows[19]: since Ω_2 increases by unity each period

$$\begin{aligned} 1 = \oint d\Omega_2 &= \oint (\partial \Omega_2/\partial q_2) \, dq_2 \\ &= \oint (\partial^2 W/\partial J_2 \partial q_2) \, dq_2 \\ &= (\partial/\partial J_2) \oint (\partial W/\partial q_2) \, dq_2 \\ &= (\partial/\partial J_2) \oint p_2 \, dq_2 \end{aligned} \tag{110}$$

The last condition is satisfied if (109) holds, but more generally it only requires J_2 to have the form

$$J_2 = \oint p_2 \, dq_2 + K(q_1, p_1) \tag{111}$$

The function K is arbitrary and varying the choice of this function (with H remaining constant) results in different choices of the function F in (102) and in the equations that follow there.

However, if we wish to use the transformation (103) as the starting point for a perturbation expansion there exists an additional consideration which dictates the choice of J_2 satisfying (109) : unless such a choice is made, the transformation of the subsidiary variables (p_1, q_1) for finite ε will contain a secular term. This point will be discussed in the next section; the function F to be used is than implicitely defined by the relation

$$J_2 = (2m)^{\frac{1}{2}} \oint \left[J_2/T + F - J^*\omega/2\pi \right]^{\frac{1}{2}} \left[dq_2/(\nabla\sigma)^2 \right] \tag{112}$$

THE SECOND PERIODICITY — FINITE ε

To transform (90) with (p_1, q_1) (but not J^*) restored to canonical status, equation (103) must be modified to

$$W(J_2, q_2, P_1, q_1) = W_0 + P_1 q_1 \tag{113}$$

where (P_1, Q_1) are new variables (though older notation is used in labeling them) and W_0 is the expression on the right side of (103), evaluated at $(cP_1/e , q_1 , q_2)$.

The transformation of (p_2, q_2) resembles the one previously obtained — in particular, (104) is again obtained. Because T, F and ω are now evaluated in "mixed" variables, this confirms the form of the zero-order Hamiltonian

$$H = J_2/T(P_1, Q_1, J^*) + F(P_1, Q_1, J^*) + O(\varepsilon) \qquad (114)$$

only if the transformation of the subsidiary variables (or at least of q_1) is a near-identical one. For q_1 we have, by (113)

$$Q_1 = q_1 + \partial W_0/\partial P_1$$

The second term here contains $\partial/\partial P_1$ which lowers the term by one order in ε — P_1 being slow — but it also contains W_0 which is an open-ended integral in q_2. Unless this term is either periodic or independent of P_1, it may grow without limit and after $O(\varepsilon^{-1})$ periods this growth may be large enough to offset $\partial/\partial P_1$ and give a transformation which no longer is of near-identity type.

Let $q_2(k)$ denote the value of q_2 at the end of the k-th period and let N be a large integer, of order ε^{-1}. Suppose the system is in its $(N + 1)$ period: then

$$\partial W_0/\partial P_1 = (\partial/\partial P_1)\left\{\sum_{k=1}^{N}\left(\int_{q_2(k-1)}^{q_2(k)} p_2\, dq_2\right) + \int_{q_2(N)}^{q_2} p_2\, dq_2\right\}$$

$$= (\partial/\partial P_1)\left\{\sum\left[J_2 + K(q_1, P_1)\right] + O(J_2)\right\} \qquad (115)$$

The last term in (115) is O(1) and differentiation reduces it to $O(\varepsilon)$, allowing it to be ignored. The summation term on the other hand is large: it consists of two parts, the first of which does not survive differentiation, since it is equal to NJ_2. The second part could raise (115) to O(1), but if J_2 is defined as in (109) K vanishes and this part does not arise.

It may be remarked here that the same consideration is implicit in the earlier transformation $(P_3, Q_3) \rightarrow (J, \Omega)$, generated by (46). In that case, again, expressions such as $\partial W/\partial p_1$ appearing in the transformation of subsidiary variables could give rise to secularly increasing terms if the open-ended integral $\partial W/\partial \delta$ exhibited such growth. As is shown in (56) this integral is in fact periodic in Ω and such growth is avoided.

The first-order Hamiltonian is

$$H = J_2 / T + F + H^{(1)} + O(\varepsilon^2) \tag{116}$$

where $H^{(1)}$ is obtained from (100) by replacing (p_1, q_1) with (P_1, Q_1) and eliminating (p_2, q_2) by means of (104) and (105). The replacement of q_2 by Ω_2 is rather difficult, but fortunately it may be by-passed in evaluating the first-order adiabatic invariant. By (11) and (8), this invariant is (ε not written, averaging taken over Ω_2)

$$J_2^* = J_2 + T\,(H^{(1)} - \langle H^{(1)} \rangle) + O(\varepsilon^2) \tag{117}$$

Let q_2 be retained in $H^{(1)}$ as an auxiliary variable replacing Ω_2. To find $\langle H^{(1)} \rangle$ one introduces q_2 as integration variable instead of Ω_2, using (105) (other variables may be ignored to lowest order):

$$\begin{aligned} \langle H^{(1)} \rangle &= \int_0^1 H^{(1)}\, d\Omega_2 = \oint H^{(1)} (\partial \Omega_2/\partial q_2)\, dq_2 \\ &= (m/2)^{\frac{1}{2}} T^{-1} \oint H^{(1)} \left[J_2/T + F - J^*\omega/2\pi\right]^{-\frac{1}{2}} \left[dq_2/(\nabla\sigma)^2\right] \end{aligned} \tag{118}$$

No attempt has been made to compare (117) and (118) to the first-order derivation of J^* presented elsewhere[20].

THE THIRD PERIODICITY

Assume now that the transformation eliminating Ω_2 has been carried out, so that the new Hamiltonian contains J_2^* as a (constant) canonical variable describing the longitudinal motion. In this transformation P_1 and Q_1 also undergo near-identity change and their transformed versions will be simply denoted P and Q, subscripts no longer being necessary.

The new Hamiltonian is

$$H(P, Q, J_2^*, J^*) = J_2^* / T(P, Q, J^*) + F + \langle H^{(1)} \rangle + O(\varepsilon^2) \tag{119}$$

where $\langle H^{(1)} \rangle$ and F are the same as in (116) and (118) but with new variables replacing old ones, a substitution which affects only higher orders since this again is a near-identity transformation. Again this may be used, via Liouville's theorem, to derive a kinetic equation for a collisionless plasma trapped in a mirror-like geometry, with averaging over both gyration and longitudinal motion. It may be mentioned here that the "Hamiltonian properties" of J_2^* / T were already noted by Northrop [15][21].

The Hamilton-Jacobi equation of (119) now seeks the generating function $W(p, Q)$ of a canonical transformation to a new set (p, q) in which p is constant and its conjugate q is linear in time. This function satisfies

$$H\left[(\partial W/\partial Q),\ Q,\ J_2^*,\ J^*\right] = \nu p \tag{120}$$

with ν a yet undetermined constant which could involve J_2^* and J^*. Assuming that $\partial W/\partial Q$ can be extracted from this

$$\partial W/\partial Q = f(\nu p, Q, J_2^*, J^*) \tag{121}$$

the equation is easily solved by quadrature.

The physical significance of this somewhat formal procedure hinges on the fact that independently of the form of W, a canonical transformation generated by it satisfies

$$\partial(p, q)/\partial(P, Q) = 1 \tag{122}$$

This is another way of saying that a canonical transformation represents an area-preserving mapping in the (p,q) phase plane, a fact already noted by Gardner[7] and by many others. One may prove (122) by direct substitution.

Now P and Q themselves are related by near-identity transformations to the variables P_1 and Q_1 of (35) which, apart from an unimportant factor (c/e), carry the connotation of guiding center Euler potentials. These potentials, however, are far from unique: if (α, β) represents one choice for them, equivalent choices $\alpha'(\alpha, \beta)$ and $\beta'(\alpha, \beta)$ may be used to describe the same magnetic field, provided

$$\partial(\alpha', \beta')/\partial(\alpha, \beta) = 1 \tag{123}$$

In view of (122) the new variables (p, q) may be identified as describing (to lowest order) a different choice of (α, β) for the field in which the particle moves. One may say that among the many different choices of (α, β) describing the given field, there exists one choice

— or rather, a set of choices — such that a given particle with specified J^* and J_2^* , on the average (average over Ω and Ω_2), stays on the same α and advances uniformly in β . The Hamilton-Jacobi equation (120) then provides a method for deriving Euler potentials belonging to this particular set.

If the new β — that is, q — is an angle-like variable, the motion will again be periodic. This will always occur in axisymmetrical fields (if the other two periodicities exist) for if then β is initially chosen as the azimuth angle φ, it will be absent from the Hamiltonian in all of its transformations. Thus in this case Q does not appear in (119) and the last transformation is trivial. However, the field needs not be axisymmetrical for the third periodicity to occur, as is demonstrated by the motion of radiation belt particles in the earth's magnetic field, which may be viewed as a distorted dipole field.

If the motion is periodic it will possess an action variable

$$J_3 = \oint p\, dq \tag{124}$$

This is commonly known as the third invariant or the flux invariant[22], since if p and q are viewed as averaged Euler potentials, then

$$J_3 = (e/c) \oint \alpha\, d\beta \tag{125}$$

and the integral appearing here equals the magnetic flux Φ embraced by the particle's orbit in one period. From general theory [6][19] one may expect it to be adiabatically conserved under slow perturbations. Our field cannot provide such perturbations, since all variables have already been accounted for; it may be shown, however[22], that J_3 will be adiabatically conserved if the field varies slowly in time or if an electric field is slowly applied to it.

APPENDIX

We wish to relate the first-order correction of (88) to geometrical quantities. Extracting a factor B from the first term and using (93) gives

$$\underline{B}\,\underline{\rho} : \nabla\nabla s_0 = \underline{\rho}\,\underline{B} : \nabla\nabla s_0$$

$$= \underline{\rho}\,\underline{B} : (\nabla\hat{\underline{B}} - \nabla\alpha\,\nabla a - \nabla\beta\,\nabla b) + O(\varepsilon^2)$$

$$= \underline{\rho}\,\underline{B} : \nabla\hat{\underline{B}} + O(\varepsilon^2) \qquad \text{(A-1)}$$

Forming the dot product of $\underline{\rho}$ with

$$\nabla \times (\underline{B} \times \hat{\underline{B}}) = 0$$

$$= \hat{\underline{B}}\cdot\nabla\underline{B} - \underline{B}\cdot\nabla\hat{\underline{B}} + \underline{B}\,(\nabla\cdot\hat{\underline{B}}) \qquad \text{(A-2)}$$

and using (93) once more gives, finally,

$$\underline{B}\,\underline{\rho} : \nabla\nabla s_0 = \underline{\rho}\,\hat{\underline{B}} : \nabla\underline{B}$$

$$= \hat{\underline{B}}\,\underline{\rho} : \nabla\underline{B} - \underline{v}_\perp^{(0)}\cdot(\nabla\times\underline{B})/\omega \qquad \text{(A-3)}$$

which, apart from non-geometrical factors, equals the $v_{/\!/}^{\,2}$ term given by Northrop[15].

Before evaluating the remaining terms, all of which are proportional to p_2 (i.e. to $v_{/\!/}$), it is useful to derive $\nabla\times\underline{v}_\perp^{(0)}$, in the system of variables that includes (J, Ω). We have, by (31) and (86)

$$m\,(\nabla \times \underline{v}_{\perp}^{(0)}) \;=\; \nabla \times \left[P_3 \nabla\alpha \;-\; (e/c)(Q_3 + \lambda\, P_3)\,\nabla\beta\right]$$

$$=\; (\nabla P_3 \times \nabla\alpha\,) \;-\; (e/c)(\nabla Q_3 + \lambda\nabla P_3 + P_3\nabla\lambda\,) \times \nabla\beta$$

$$=\; (P_3\,/2\delta\,)(\nabla\delta \times \underline{T}) \;+\; (e/c)(Q_3\,/2\delta\,)(\nabla\delta \times \nabla\beta\,) \;-\; (e/c)P_3(\nabla\lambda \times \nabla\beta\,) \qquad \text{(A-4)}$$

Using (32) and (97)

$$(m^2/B)\;\underline{v}_{\perp}^{(0)}\cdot(\nabla \times \underline{v}_{\perp}^{(0)}) \;=\; (m/B)\;\left[P_3\,\underline{T} - (e/c)Q_3\nabla\beta\right]\cdot(\nabla \times \underline{v}_{\perp}^{(0)})$$

$$=\; -\,(Q_3\,P_3\,/\,\delta\,)(\hat{\underline{B}}\cdot\nabla\delta) \;+\; (e/c)\;P_3^{\,2}\;(\hat{\underline{B}}\cdot\nabla\lambda) \qquad \text{(A-5)}$$

By (41)

$$(\hat{\underline{B}}\cdot\nabla\delta\,) \;=\; \partial\delta/\partial\sigma \;+\; o(\varepsilon^2) \qquad \text{(A-6)}$$

Comparing this to (84), (77) and (56) shows that the $0(\varepsilon)$ correction term proportional to $v_{/\!/}$ in J^* is

$$(2\pi/\omega)\;p_2\left\{\,\hat{\underline{B}}\;\underline{\rho}\;:\;\nabla\,\underline{v}_{\perp}^{(0)} \;-\; (1/2\omega)\;\underline{v}_{\perp}^{(0)}\cdot(\nabla \times \underline{v}_{\perp}^{(0)})\right\}_{osc} \qquad \text{(A-7)}$$

Replacing the factor $\underline{v}_{\perp}^{(0)}$ by $\omega\,(\underline{\rho} \times \hat{\underline{B}})$ and applying (93) gives for this term

$$(\pi/\omega)\;p_2\left\{\,\hat{\underline{B}}\,\underline{\rho}\;:\nabla\,\underline{v}_{\perp}^{(0)} \;+\; \underline{\rho}\;\hat{\underline{B}}\;:\nabla\;\underline{v}_{\perp}^{(0)}\right\}_{osc} \qquad \text{(A-8)}$$

The average of the first term in (A-8) is given by (98) and is easily subtracted: the result agrees with Northrop's $v_{/\!/}$ term. The second term remains unaccounted for and has not been further evaluated.

REFERENCES

(1) H. Poincaré, Les Methodes Nouvelles de la Mecanique Celeste, Vol. II, Gauthier-Villars, Paris 1893; reprinted by Dover Publications, New York, 1957; NASA translation TT F-451, 1967

(2) H. Von Zeipel, Ark. Astr. Mat. Fys. 11, no. 1 (1916)

(3) H.C. Corben and P. Stehle, Classical Mechanics, 2nd Ed. ; John Wiley and sons, 1960 ; Sect. 75 .

(4) D. Ter Haar, Elements of Hamiltonian Mechanics, North Holland Publishing Co., 1961 ; Chapter 7 .

(5) P. Musen, Journal of the Astronautical Sciences 12, 129 (1965)

(6) D.P.Stern, Classical Adiabatic Perturbation Theory, J. Math. Phys. 1971 (to be published).

(7) C. Gardner, Phys. Rev. 115, 791 (1959)

(8) G. Contopoulos, J. Math. Phys. 7, 788, (1966)

(9) D.P. Stern, J. Math. Phys. 11, 2776 (1970)

(10) D.P. Stern, Celestial Mechanics 3, 241 (1971)

(11) D.P. Stern, Am. J. Phys. 38, 494 (1970) and references cited there.

(12) H. Goldstein, Classical Mechanics, Addison-Wesley 1953 ; Section 9-2 .

(13) Ref. 3 , section 62 .

(14) H. Grad, in Proc. 18th Symp. Appl. Math. 1967, p. 162 (1967)

(15) T.G. Northrop, The Adiabatic Motion of Charged Particles (John Wiley and Sons, Inc., New York, 1963)

(16) J.G. Siambis and T.G. Northrop, Phys. Fluids 9, 2001 (1966) ; eq. 6 .

(17) L. Brand, Vector and Tensor Analysis (John Wiley and Sons, Inc., New York, 1947); section 85, eq. 9 .

(18) G.F.Chew, M.L. Goldberger and F.E. Low, Proc. Roy. Soc. A236, 112 (1956)

(19) M. Born, The Mechanics of the Atom, 1925, 2nd Edit. of English Translation published by Frederick Ungar Publ. Co., 1960 ; Section 10 .

(20) T.G. Northrop, C.S. Liu and M.D. Kruskal, Phys. Fluids 9 , 1503 (1966)

(21) T.G. Northrop, Reviews of Geophysics 1, 283 (1963)

(22) T.G. Northrop and E. Teller, Phys. Rev. 117, 215 (1960)